ADVANCES IN MATHEMATICS 53, 28–54 (1984)

The Wronskian Formalism for Linear Differential Equations and Padé Approximations

D. V. CHUDNOVSKY* AND G. V. CHUDNOVSKY*

Department of Mathematics, Columbia University, New York, New York 10027

0. INTRODUCTION

This paper is devoted to rational approximations. While rational approximations to numbers in various archimedian and nonarchimedian metrics constitute the subject of diophantine approximations, the rational approximations to functions, that we study here, are known in analysis as Padé approximations. The immediate purpose of the paper is the proof of a few general theorems on the determination of the best rational approximations to solutions of linear differential equations with rational function coefficients. However, the analytical methods used in the proof are more important from the author's point of view, since they provide an analytic tool for the solution of many problems from diophantine approximations to numbers (and functions alike).

The functional problem on the determination of the best rational approximations (Padé approximations) is analogous to the diophantine problem of rational aproximations to algebraic numbers, studied extensively since Liouville (1844). For a concise exposition and references see Baker [4]. The high points in the solution of this diophantine problem are the Roth theorem [32] and the Schmidt theorem [34], which furnish the correct order of approximations, or simultaneous approximations, of algebraic numbers by rationals. For algebraic functions the results of Roth's theorem form (approximations by rational functions in nonarchimedian metrics) were proved by Uchiyama [36] (cf. Mahler [24]). The analog of the Schmidt theorem for simultaneous rational approximations of algebraic functions was proved by Ratliff [30]. These results, as well as the original theorems of Roth and Schmidt, are noneffective in the sense that constants in the measure of approximations cannot be determined effectively from the proof.

In the functional case, the natural class of functions, for which the problem of rational approximations should be considered, is given by the

*This work was supported in part by John Simon Guggenheim Memorial Foundation, and U.S. Air Force Grant AFOSR-81-0190.

0001-8708/84 $7.50

solutions of differential equations. Kolchin [18], in 1959, put forward a conjecture on the best possible rational approximations of the Roth's form. Kolchin's problem deals with the finite differential extensions of the field $k(x)$ of functions rational over an algebraically closed field k of characteristic zero. Let $f(x)$ be a solution of an algebraic differential equation over $k(x)$ defined by its formal power series expansion in the neighborhood of a point x_0. For every $\varepsilon > 0$ Kolchin's conjecture asserts the existence of a constant c depending on $f(x)$, ε and x_0 such that for the arbitrary polynomials $P(x)$ and $Q(x)$ from $k[x]$, we have

$$\operatorname{ord}_{x=x_0}\left(f(x)-\frac{P(x)}{Q(x)}\right) \leqslant (2+\varepsilon)\cdot\max\{\deg(P),\deg(Q)\}+c.$$

Entirely new and effective methods for the solution of Kolchin's problem for algebraic differential equations were proposed by Osgood [26]. His work [25–28] and Schmidt's results, see [33], moved forward significantly the effectivization of the Thue–Siegel–Roth theorem for algebraic functions, with the best possible results in some cases.

In this paper we prove an effective solution to Kolchin's problem and prove an effective form of the Schmidt theorem on the best simultaneous rational approximations, in the case of solutions of arbitrary linear differential equations with rational function coefficients.

Our methods of proof that use Wronskians and Wronskian-type invariants serve as a natural analytic basis for the study of rational approximations to functions and values of the functions.

The methods of construction of auxiliary nonlinear differential equations by means of Wronskians of solutions of linear differential equations, that are used in this paper, are natural in the theory of completely integrable systems [3, 5], as studied by the authors in [7–9]. The typical feature of these methods is the use of the Riccati-type equations satisfied by ratios of solutions of the auxiliary linear differential equations (see, particularly Chaps. 1 and 4). Another important feature of our methods is the use of symmetric or antisymmetric parts of the tensor products of spaces of solutions of linear differential equations. (See particularly Chap. 5.) The relation of these methods with invariant theory (cf. [37]) will be exploited further.

Our functional results are, in fact, corollaries of our deeper number-theoretic results from which the present functional versions grew up. In the functional case, the solution of Kolchin's problem, is equivalent to the normality and the almost normality (or perfectness and almost perfectness) of Padé approximations (or Hermite–Padé approximations in the case of several functions). This point will be expanded upon in further papers in connection with the conjecture that $\varepsilon = 0$ in Kolchin's problem or the functional versions of the theorems of Roth and Schmidt. It should be pointed out, however, that the existing analytic methods of proving the normality of Padé approximations depend heavily on the positivity properties. The best result in this direction belongs to Arms and Edrei [1, 13], where functions such as $(\cos x)^k$, $\tan x/x,\ldots$, etc. were considered.

We want to note that our conjecture that $\varepsilon = 0$ in Kolchin's problem is the feature of the rational approximation problem in the functional case only. For numbers it seems highly implausible that one can have $\varepsilon = 0$.

While the results of this paper deal with linear differential equations only, our methods allow us to establish the solution of Kolchin's problem for arbitrary differential equations.

Our paper is organized as follows. In Chapter 1 the notion of a Bäcklund transformation and its connection with Wronskian formulas is described. Bäcklund transformations and related Ricatti equations, that play an important role in the theory of completely integrable systems [3], are the main auxiliary instruments in the proofs below. Other auxiliary results, from the theory of Picard–Vessiot extensions of differential fields, generated by linear differential equations, are summarized in Chapter 2. The main purpose of this paper is to establish effective functional analogs of the Schmidt theorem for arbitrary solutions of linear differential equations with rational function coefficients. The most general result in this direction is presented in Chapter 3, where the functional approximation theorem provides the effective Schmidt theorem for simultaneous approximations in several nonarchimedian metrics. In the case of algebraic functions this result provides an effective version of the Ratliff theorem [30] and the generalization of Mahler's approximation theorems [24]. We also want to point out the discussion following the proof of the functional approximation theorem of Chapter 3, where the estimates are sharpened with an error term depending on the dimension of the graded ring generated by solutions of linear differential equations. We conjecture, however, that one can put $\varepsilon = 0$ in the functional approximation theorem, which is equivalent to the constant error term (and the almost normality of Padé approximations). The arguments of Chapter 3 are specialized in Chapter 4 in the case of approximations of a solution of a linear differential equation by rational functions (the effective Roth Theorem 4.5). The appearance of the auxiliary nonlinear differential equations, their connection with the Ricatti equations and the importance of Kolchin's concept of differential denomination [18, 26], becomes clear in the proof of Theorem 4.5. In Chapter 5 we use the methods of Bäcklund transformations to establish $\varepsilon = 0$ in Schmidt's theorem for a class of algebraic functions. Our result, Theorem 5.1 generalizes Osgood's result [25] proved for cubic algebraic functions.

This paper is the first in a series of related articles, in which the authors apply new analytic methods, indicated in this paper, to the problems of number theory. In the sequel we investigate in more detail graded subrings of differential fields and generalized Wronskians associated with them. This implys the solution of Kolchin's problem for algebraic differential equations. Particular attention will be devoted to the effectivization of all constants in the positive solution of the Kolchin problem. Classes of differential equations, for which we can prove the conjecture $\varepsilon = 0$ of the present paper, will be examined as well.

Functional results are specializations of more general number-theoretical results. The present paper serves as an analytic introduction to our further papers, where the Roth and Schmidt theorems are proved for values of functions satisfying linear differential equations. While the analytic part of the proof uses the general functional methods of this paper, sharp number-theoretic results require additional assumptions on the arithmetic properties of linear differential equations. From the point of view of these arithmetic properties, two natural classes of functions are those of the E- and G-functions introduced by Siegel [35]. In the short exposition below, we consider linear differential equations over $\mathbb{Q}(x)$.

Following Siegel [35], a solution $f(x)$ of a linear differential equation with coefficients from $\mathbb{Q}(x)$ is called an E-function, if at $x = 0$ it has a Taylor expanion of the form

$$f(x) = \sum_{n=0}^{\infty} \frac{a_n}{n!} x^n,$$

where $a_n \in \mathbb{Q}$, and for every $\varepsilon > 0$ we have $|a_n| \leqslant n!^{\varepsilon}$ and the common denominator of $a_0, ..., a_n$ is bounded by $n!^{\varepsilon}$ for $n \geqslant n_0(\varepsilon)$.

Similarly, one calls a solution $f(x)$ of a linear differential equation with coefficients from $\mathbb{Q}(x)$ a G-function, if its Taylor expansion at $x = 0$ has the form

$$f(x) = \sum_{n=0}^{\infty} a_n x^n,$$

where $a_n \in \mathbb{Q}$, and both a_n and the common denominator of $a_0, ..., a_n$ grow not faster than a geometric progression of n.

The class of E-functions includes e^x, the Bessel functions $J_\nu(x)$ for $\nu \in \mathbb{Q}$, and the entire hypergeometric functions ${}_qF_p(x|{}^{a_1,...,a_q}_{b_1,...,b_p})$ for $q \leqslant p$ and rational numbers a_i, b_j. Among the examples of G-functions one finds algebraic functions regular at $x = 0$, Abelian integrals and generalized hypergeometric functions ${}_{p+1}F_p(x|{}^{a_1,...,a_{p+1}}_{b_1,...,b_p})$ for rational numbers a_i, b_j. Classes of E- and G-functions admit wide effective versions of the Roth and Schmidt theorems in the direction predicted by the conjectures of Lang [22]. In order to formulate our number-theoretic results, and to compare them to the functional results of this paper, we introduce the following diophantine property:

DEFINITION 0.1. A number field $\mathbb{K}$ has a property (S), if for the arbitrary elements $\theta_1,..., \theta_n$ from $\mathbb{K}$ and any $\varepsilon > 0$ there exists a constant $c = c(\theta_1,..., \theta_n, \varepsilon) > 0$ with the following property: For the arbitrary rational integers $H_1,..., H_n$ and $H = \max(|H_1|,..., |H_n|)$ we have

$$|H_1\theta_1 + \cdots + H_n\theta_n| > H^{-n+1-\varepsilon}$$

provided that $H \geqslant c$ and $H_1\theta_1 + \cdots + H_n\theta_n \neq 0$.

The property (S) is a reformulation of the Schmidt theorem for arbitrary elements of the field $\mathbb{K}$. Until now the only class of examples of number fields having a property (S) was provided by the Schmidt theorem that $\bar{\mathbb{Q}}$ has a property (S) [34]. We want to point out that it is still an open question, whether for almost all θ the field $\mathbb{Q}(\theta)$ has a property (S).

A property (S) can be formulated for fields with nonarchimedian normings and functional fields, as a substitute for the notion of the Schmidt theorem, which we use everywhere in this paper. In this case the results of the present paper show that arbitrary subfields of Picard–Vessiot (and composite Picard–Vessiot) extensions of $k(x)$ have a property (S) for an algebraically closed field of constants k of characteristic zero.

Our main number-theoretical results, completely proof of which will appear later in this series of papers, show that the addition of values of E-functions at rational points to $\mathbb{Q}$ generates fields with property (S). Similar results under some conditions hold for values of G-functions as well. We present first the result for values of E-functions:

THEOREM 0.2. *Let $f_1(x),..., f_n(x)$ be arbitrary E-functions and let $r \in \mathbb{Q}$, $r \neq 0$. Then for any $\varepsilon > 0$ there exists an effective constant $c_1 = c_1(f_1,..., f_n, r, \varepsilon) > 0$ such that for the arbitrary rational integers $H_1,..., H_n$ we have*

$$|H_1 f_1(r) + \cdots + H_n f_n(r)| > H^{-n+1-\varepsilon}$$

for $H = \max(|H_1|,..., |H_n|)$ provided that $H \geqslant c_1$ and $H_1 f_1(r) + \cdots + H_n f_n(r) \neq 0$.

This results answer some of the questions of Lang [22]. Theorem 0.2 generalizes the previous results of [10], where it was proved that the addition of values of exponential functions at rational points to $\mathbb{Q}$ generates fields with property (S).

Results similar to Theorem 0.2 for G-functions require the additional (G, C)-function condition described in [11]:

THEOREM 0.3. *Let $f_1(x),..., f_n(x)$ be (G, C)-functions, linearly independent over $\mathbb{C}(x)$, and let $r = a/b$ for a and b being rational integers such* *that* $|b| > |a|^n$. *For* $\varepsilon > 0$ *there exists an effective constant* $c_2 = c_2(f_1, ..., f_n, r, \varepsilon) > 0$ *such that the following property is satisfied. For the arbitrary rational integers* $H_1, ..., H_n$ *and* $H = \max(|H_1|, ..., |H_n|)$ *we have*

$$|H_1 f_1(r) + \cdots + H_n f_n(r)| > H^{\mu - \varepsilon}$$

for $\mu = -(n-1)\log|b|/\log|b/a^n|$, *provided that* $H \geqslant c_2$ *and* $|b| \geqslant B_0(f_1, ..., f_n, \varepsilon)$.

Theorem 0.3 can be applied to the algebraic functions $f_1(x), ..., f_n(x)$, when r is sufficiently close to 0. This provides an effectivization of the Roth theorem and the Schmidt theorem for a large class of algebraic numbers defined as values of algebraic functions given by a Taylor series in the neighborhood of their regular points. Such a result gives, in particular, a far-reaching effective improvement over the Liouville theorem for many interesting diophantine problems.

The authors want to acknowledge the importance of Osgood's contribution in the field of rational approximations. We want to point out Osgood's announcement that he proved the theorems of Roth and Schmidt for algebraic functions. Moreover, according to his recent communication, Osgood proved the Schmidt theorem for solutions of linear differential equations as well.

We want to thank Kolchin for his illuminating remarks and discussions of the proofs. The authors derived important insights on the structure of the auxiliary differential operators from the use of the formal manipulation programs SCRATCHPAD and SMP. We thank the computer algebra group of IBM for the opportunity to use SCRATCHPAD, and S. Wolfram for introducing us to SMP.

1. Wronskians, Bäcklund Transformations and Ricatti Equations

The Wronskian formalism has many things in common with classical invariant theory, at least if looked upon from the formal point of view using the concept of bideterminants, see [37], the section invariant theory in the Introduction. Such a formal identification can be continued further using the deep relationship with the K-theory, λ-rings and representations of infinite dimensional Lie algebras of the $A_n^{(1)}$-type as described in [23]. Without these formalizations, Wronskians were already widely used in the study of auto-Bäcklund transformations of completely integrable systems of "soliton" or Korteweg–de Vries (KdV) form, see [3, 5] (and the references there). It is here, through the Wronskian formalism, that the Ricatti equations and generalizations of Ricatti equations appear as nonlinear differential equations associated with linear differential (spectral) problems. The Wronskian formulas explicitly appeared for the first time in Darboux's studies of transformations of linear differential equations (of the second order) [11, 12]. The problem that Darboux posed and solved is the following one. How to describe a map (differential correspondence) that transforms a differential operator all of whose eigenfunctions are known, into a new differential operator of a similar form all of whose eigenfunctions are known and the coefficients of which are determined in terms of the old eigenfunctions. We present the corresponding result following [7–9].

We consider an arbitrary linear differential (scalar) spectral problems

$$L_m \psi = \lambda^m \psi \tag{1.1}$$

or

$$L_m \psi = \frac{\partial}{\partial t} \psi \tag{1.2}$$

for a linear differential operator L_m of order m,

$$L_m = \sum_{i=0}^{m} u_i \left(\frac{d}{dx}\right)^i, \tag{1.3}$$

with the normalization $u_m = 1$, $u_{m-1} = 0$. Then for any n linearly independent eigenfunctions $\psi_1(\lambda_1),\dots, \psi_n(\lambda_n)$ the general solution $\psi(\lambda) \overset{\text{def}}{=} \psi_0(\lambda)$ of the linear problem (1.1) or (1.1′) is mapped onto the general solution of a similar linear problem

$$\bar{L}_m \cdot \psi_{\bar{B}} = \lambda^m \cdot \psi_{\bar{B}} \tag{1.1′}$$

or

$$\bar{L}_m \cdot \psi_{\bar{B}} = \frac{\partial}{\partial t} \psi_{\bar{B}} \tag{1.2′}$$

for

$$\bar{L}_m = \sum_{i=0}^{m} \bar{u}_i \cdot \left(\frac{d}{dx}\right)^i, \qquad \bar{u}_m = 1, \bar{u}_{m-1} = 0. \tag{1.3′}$$

Here is the expression of the new eigenfunction $\psi_{\bar{B}}(\lambda)$(called the Bäcklund transformation of $\psi(\lambda)$),

$$\bar{B} : \psi(\lambda) \mapsto \psi_{\bar{B}}(\lambda) \overset{\text{def}}{=} \frac{W(\psi(\lambda), \psi_1(\lambda_1),\dots, \psi_n(\lambda_n))}{\prod_{i=1}^{n} (\lambda - \lambda_i) \cdot W(\psi_1(\lambda_1),\dots, \psi_n(\lambda_n))}, \tag{1.4}$$

where $W(f_1,\dots,f_k)$ is the Wronskian of $f_1,\dots,f_k$, $W(f_1,\dots,f_k) \stackrel{\text{def}}{=} \det((d/dx)^{j-1}f_i)_{i,j=1,\dots,k}$. The Bäcklund transformation (1.4) transforms the coefficients u_i of L_m into $\bar{u}_i$ of $\bar{L}_m$ following simple rules involving the derivatives of Wronskian of $\psi_1(\lambda_1),\dots,\psi_n(\lambda_n)$. For example, $\bar{B}: u_{m-2} \mapsto \bar{u}_{m-2} \stackrel{\text{def}}{=} u_{m-2} + m(d/dx)^2 \log W(\psi_1(\lambda_1),\dots,\psi_n(\lambda_n))$. In general, coefficients of $\bar{L}_m$ are expressed in terms of the transformations of the single "pseudo-potential" σ defined as $u_{m-2} = m(d^2/dx^2)\log\sigma$, with the following action of the Bäcklund transformation

$$\sigma_{\bar{B}} = \sigma \frac{\det((d/dx)^{j-1}\,\psi_i(\lambda_i))_{i,j=1,\dots,n}}{\det(\lambda_i^{j-1})_{k,j=1,\dots,n}}.$$

Though the Darboux formulas were derived in 1882, their discrete versions were derived by Christoffel in 1858 [6]. The relationship of the Christoffel formula with the Darboux expression of Bäcklund transformations and their spectral sense were pointed out by Krein [20].

The Ricatti equation associated with the linear problem (1.1) or (1.2) arises as an elementary Bäcklund transformation, $n=1$, in the constructions above. The Ricatti equation in this context can be described as a nonlinear differential equation of order $m-1$ satisfied by ψ'/ψ for a general eigenfunction ψ of (1.1) or (1.2). For completely integrable systems of KdV type the use of the auxiliary Ricatti equation in the description of Bäcklund transformations is known as the Miura transformation [3]. For matrix generalizations of the KdV system one has to consider several linear differential equations of the form (1.1) or (1.2). Hence we define a generalized Ricatti equation associated with the two linear differential equations $L_1[f_1]=0$ and $L_2[f_2]=0$ as a nonlinear differential equation whose general solution is f_1/f_2. This nonlinear differential equation can be explicitly represented in terms of Wronskians of fundamental systems of solutions of $L_1=0$ and $L_2=0$. Namely, let $h_1,\dots,h_m$ and $g_1,\dots,g_l$ be fundamental systems of solutions of the linear differential equations $L_1[y]=0$ and $L_2[y]=0$, respectively. Then the generalized Ricatti equation associated with L_1 and L_2 has the following explicit representation:

$$\mathfrak{D}_{12}[Y]=0$$

for

$$\mathfrak{D}_{12}[Y] \stackrel{\text{def}}{=} \frac{W(Y\cdot g_1,\dots,Y\cdot g_l,h_1,\dots,h_m)}{W(g_1,\dots,g_l)\cdot W(h_1,\dots,h_m)}.$$

The coefficients of the nonlinear differential operator $\mathfrak{D}_{12}[Y]$ are elements of the differential closure of the field containing all the coefficients of L_1 and L_2. The fact that $\mathfrak{D}_{12}[Y]=0$ is a representation of the generalized Ricatti equation follows immediately from the main property of Wronskians:

LEMMA 1.1. *The Wronskian $W(f_1,\dots,f_m)$ of functions $f_1(x),\dots,f_m(x)$ from the differential field with the field of constants k is identically zero if and only if $f_1,\dots,f_m$ are linearly dependent over k.*

2. AUXILIARY RESULTS ON PICARD–VESSIOT EXTENSIONS

We use here the standard terminology of differential algebra and the results of the Picard–Vessiot theory of linear differential equations [19].

We start with an algebraically closed field of constants k and a differential field K over k with a differentiation $d/dx = {}'$. The most interesting example for us is $K = k(x)$ and $k = \mathbb{C}$.

The Picard–Vessiot extension of K is characterized as a differential extension of K having the same field of constants k, and generated by the fundamental system of solutions of a linear differential equation with coefficients from K. This means that we start from a homogeneous linear differential equation

$$L[y] \overset{\text{def}}{=} y^{(n)} + a_{n-1} y^{(n-1)} + \cdots + a_1 y' + a_0 y = 0,$$

with coefficients from the field K. We choose n solutions $y_1,\dots,y_n$ of $L[y] = 0$, linearly independent over the field of constants k. The differential field generated by the addition of $y_1,\dots,y_n$ to K, denoted by $M = K\langle y_1,\dots,y_n\rangle$, is a Picard–Vessiot extension of K having the same field of constants k. The field M can be described as a field of fractions of the ring $K\{y_1,\dots,y_n\}$ of differential polynomials in $y_1,\dots,y_n$ over K.

The group G of differential isomorphisms of M, the Picard–Vessiot group of M, is an algebraic matrix group. The action of any differential isomorphism σ of M from G is linear on the generators y_i of M. This means that with $\sigma \in G$ we associate a nonsingular $n \times n$ matrix $m = (m_{ij})$ from $GL(n; k)$, such that $y_i\sigma = \sum_{j=1}^n m_{ij} y_j$ (m_{ij} are constants from k), $i = 1,\dots, n$.

According to the Picard–Vessiot theory, the field M is a normal extension of K (strongly normal in the sense of Kolchin [19]). This means that an element of M, which is invariant under the action of G, belongs to K. The normality provides an easy test to determine whether an element of $M = K\langle y_1,\dots,y_n\rangle$ belongs to K.

For the proofs that follow we need a description of some particular graded subrings of $K\{y_1,\dots,y_n\}$. We start with the ring $k[y_1,\dots,y_n]$ of polynomials in $y_1,\dots,y_n$ with coefficients from k. This ring has a natural grading according to the degree of monomials, $k[y_1,\dots,y_n] = \sum_{N\geqslant 0} \mathfrak{M}_N$, where $\mathfrak{M}_N$ consists of homogeneous polynomials in $y_1,\dots,y_n$ of degree N. The ring $k[y_1,\dots,y_n]$ can be identified with an algebraic subvariety V of $k\mathbb{P}^{n-1}$. For this we consider a natural imbedding $k[x_1,\dots,x_n] \to k[y_1,\dots,y_n]$ of the ring of polynomials over k in the undetermined $x_1,..., x_n$. This defines a homogeneous ideal $I = I(V)$ in $k[x_1,..., x_n]$ generated by homogeneous polynomials from $k[x_1,..., x_n]$ that are annihilated at $y_1,..., y_n$. Hence one can consider $k[y_1,..., y_n]$ to be a finitely generated $k[x_1,..., x_n]$-module $k[x_1,..., x_n]/I$. Thus, according to the Serre–Hilbert theorem, the dimension $\dim_k \mathfrak{M}_N$ of $\mathfrak{M}_N$ over k is an integer-valued polynomial $P(N)$ for $N \geqslant N_0$ [2].

For every $N \geqslant 0$ we choose a basis $f_i^{(N)}: i \in M_N$ of the module $\mathfrak{M}_N$ over k, $\mathrm{Card}(M_N) = \dim_k \mathfrak{M}_N$ (i.e., $f_i^{(N)}$ can be chosen as monomials in $y_1,..., y_n$).

The Picard–Vessiot group G of M acts linearly on $\mathfrak{M}_N$ turning $\mathfrak{M}_N$ into a G-module. According to the description of the action of G on M above, the action of an isomorphism $\sigma \in G$ on generators $f_i^{(N)}$ of $\mathfrak{M}_N$ is represented by a nonsingular $\mathrm{Card}(M_N) \times \mathrm{Card}(M_N)$ matrix. Namely, for some $\bar{m}^{(N)} = (m_{ij}^{(N)})$ from $GL(\mathrm{Card}(M_N); k)$ we have $f_i^{(N)}\sigma = \sum_{j \in M_N} m_{ij}^{(N)} f_j^{(N)}$. This action of G on $\mathfrak{M}_N$ is naturally extended to the differential field $M^{(N)}$ containing $\mathfrak{M}_N$ (another Picard–Vessiot extension of K), $\mathfrak{M}_N \subset M^{(N)} \subset M$.

Results on the approximations in the function field case are formulated in terms of nonarchimedian mornings and valuations of differential fields. Following [18], the nonarchimedian norming $|\cdot|$ on a differential field K is assumed to satisfy the following conditions: there exist elements α and β of the value group of the norm function such that

$$\alpha\,|a| \leqslant |a'| \leqslant \beta\,|a|$$

for every element a of the field K with $|a| < 1$. For example, on the field of constants k, the norming is trivial (cf. [18] for further properties of the normings in the differential fields).

The normings of functions is usually expressed in terms of orders (valuation function) of a function at a given divisor.

In the case $K = k(x)$ with an algebraically closed field of constants k, all nontrivial normings on K have the following form: For a given $a \in k$, we define $|(x-a)^n u(x)/v(x)|_a = \rho^{-n}$ for $u(x)$, $v(x) \in k[x]$, $u(a) \neq 0$, $v(a) \neq 0$, $n \in \mathbb{Z}$. There also exists a nonarchimedian norming $|\cdot|_\infty$ on $k(x)$ defined as follows: $|u(x)/v(x)|_\infty = \rho^{-\deg(v)+\deg(u)}$. Here $\rho > 1$ is a fixed (real) constant, usually $\rho = e$. Any norming $k(x)$, trivial on k, is equivalent to one of the $|\cdot|_a: a \in k \cup \{\infty\}$.

The norming $|\cdot|_a$ can be extended from functions rational on k to a function defined by a formal power series expansion with coefficients from k. Let $k((x-a))$ be a field of the formal power series $u = \sum_{n=m}^{\infty} u_n(x-a)^n$ for some $m, -\infty < m < \infty$. Then we put $\mathrm{ord}_a(u) = \min\{n: u_n \neq 0\}$; and $|u|_a = \rho^{-\mathrm{ord}_a(u)}$ is an extension of $|\cdot|_a$ from $k(x)$ to $k((x-a))$. In the case of $a = \infty$, one replaced a local parameter $(x-a)$ by x^{-1}.

This determines the norming of functions regular at $x = a$, or having at most a pole at $x = a$. Similarly the norming is extended to functions defined by Puiseux expansions at $x=a$; in particular, to solutions of Fuchsian linear differential equations over $K=k(x)$.

3. The Effective Schmidt Theorem and the Functional Approximation Theorem for Several Valuations

In this chapter we present the complete proof of the effectivization of Schmidt's theorem for arbitrary n functions satisfying linear differential equations over $k(x)$. It should be noted that "noneffective" Schmidt's theorem was proved only for algebraic functions, and our results constitute a solution of Schmidt's problem [33]. Moreover, we prove the general theorem on the simultaneous approximation in several normings which generalizes Mahler's approximation theorem [24] for an arbitrary number of functions. We formulate and prove this theorem by induction on the number of functions, considering arbitrary normings $|\cdot|_a$, $a \in k$ extended from $k(x)$. Simple modifications allow us to include the norming $|\cdot|_\infty$ as well (cf. the transformation $x \to x^{-1}$ of $k(x)$).

As before, we consider linear differential equations over $K=k(x)$ for an algebraically closed field k of characteristic 0.

The Functional Approximation Theorem. *Let $l_1[y],\dots,l_n[y]$ be n linear differential operators with coefficients from $K=k(x)$, and let $A_0, A_1,\dots,A_n$ be sets of (finite) points of $a\in k$ corresponding to normings $|\cdot|_a$ from $k(x)$, A_0 distinct from $A_1,\dots,A_n$. For any $\varepsilon>0$ there exist constants $c_\varepsilon^0, c_\varepsilon^1$ such that for arbitrary polynomials $P_1(x),\dots,P_n(x)$ from $k[x]$ and arbitrary solutions $f_1(x),\dots,f_n(x)$ of equations $l_1[f_1]=0,\dots,l_n[f_n]=0$, respectively, that are regular at all $x=x_\alpha$ for $x_\alpha \in A_0$, we have*

$$\prod\nolimits_{x_\alpha\in A_0} |P_1(x)f_1(x)+\cdots+P_n(x)f_n(x)|_{x_\alpha}\cdot\prod_{i=1}^{n}\prod_{a\in A_i}|P_i|_a \geqslant e^{c_\varepsilon^0}\{|P_1|_\infty\cdot\cdots\cdot|P_n|_\infty\}^{-1-\varepsilon}, \tag{3.1}$$

or

$$\sum_{x_\alpha\in A_0}\operatorname{ord}_{x=x_\alpha}(P_1(x)f_1(x)+\cdots+P_n(x)f_n(x)) + \sum_{i=1}^{n}\sum_{a\in A_i}\operatorname{ord}_{x=a}(P_i)\leqslant c_\varepsilon^0+(1+\varepsilon)\cdot\left\{\sum_{i=1}^{n}\deg(P_i)\right\}, \tag{3.2}$$

provided that $\sum_{i=1}^n P_i(x)f_i(x)\not\equiv 0$ and $\sum_{i=1}^n \deg(P_i)\geqslant c_\varepsilon^1$.

Remark. Constants c_ε^0 and c_ε^1 can be easily effectively estimated in terms of the coefficients of the differential operators $l_1[y],\dots,l_n[y]$, cardinalities of sets $A_0, A_1, \ldots, A_n$ and ε. Moreover, c_ε^1 is independent of $A_0, A_1, \ldots, A_n$, and c_ε^0 depends linearly on cardinalities of $A_0, A_1, \ldots, A_n$. Constants c_ε^0 and c_ε^1 do not depend on functions $f_1(x), \ldots, f_n(x)$, and the only assumption on $f_1(x), \ldots, f_n(x)$ is their regularity at $x = x_\alpha$ for $x_\alpha \in A_0$ (or even the regularity of $R(x) = P_1(x) f_1(x) + \cdots + P_n(x) f_n(x)$ at $x = x_\alpha$). The similar results take place, if one considers functions $f_1(x), \ldots, f_n(x)$ regular at $x = \infty$.

Proof of the Functional Approximation Theorem. The proof of this theorem is by induction in n. Let us assume that the theorem is proved for any $n-1$ linear differential equations over $k(x)$. Let us prove this theorem for n linear differential equations $l_i[y] = 0$ over $k(x)$, $i = 1, \ldots, n$.

We consider the composite Picard–Vessiot extension M of the differential field $k(x)$ generated by all solutions y_i of $l_i[y_i] = 0$ for $i = 1, \ldots, n$. Then, as above in Section 2, we consider the natural grading of the ring $k[y_1, \ldots, y_n] = \sum_{N \geqslant 0} \mathfrak{M}_N$. For a given $N \geqslant 0$, we denote here by $\mathfrak{M}_N$ a vector space over k, generated by all monomials $y_1^{m_1} \cdot \cdots \cdot y_n^{m_n}$ of degree N, $m_1 + \cdots + m_n = N$, for arbitrary nonzero solutions y_i of $l_i[y_i] = 0 : i = 1, \ldots, n$.

Let us denote by $F_\gamma : \gamma \in M_N$, $\mathrm{Card}(M_N) = \dim_k \mathfrak{M}_N$, the basis of $\mathfrak{M}_N$ over k. We define then for $N \geqslant 1$ the following auxiliary differential polynomial in the differential undetermined $P_1, \ldots, P_n$,

$$\mathfrak{D}_N(P_1, \ldots, P_n) = \frac{W(\{P_i \cdot F_\gamma : i = 1, \ldots, n; \gamma \in M_N\}).}{W(\{F_\gamma : \gamma \in M_N\})^n.} \tag{3.3}$$

Here $W(h_1, \ldots, h_m) = \det((d/dx)^{l-1}(h_s))_{l,s=1,\ldots,m}$ is a Wronskian of functions $h_1 = h_1(x), \ldots, h_m = h_m(x)$. Since functions F_γ are linearly independent over k, $W(\{F_\gamma : \gamma \in M_N\}) \neq 0$.

The expression $\mathfrak{D}_N(P_1, \ldots, P_n)$ can be considered as an element of the Picard–Vessiot extension of the field $K' = K\langle P_1, \ldots, P_n \rangle$ for differentially algebraically independent $P_1, \ldots, P_n$, obtained by adding solutions y_i of $l_i[y_i] = 0$: $i = 1, \ldots, n$. The action of the Picard–Vissiot (Galois) group of this extension is linear on $\mathfrak{M}_N$ according to Section 2. Hence in (3.3), $\mathfrak{D}_N(P_1, \ldots, P_n)$ is invariant under the action of the Galois group of this Picard–Vessiot extension. From the normality of M (Picard theorem, see [19, 29]), we obtain the following:

$\mathfrak{D}_N(P_1, \ldots, P_n)$ is a differential polynomial in $P_1, \ldots, P_n$ of an order at most $\mathrm{Card}(M_N)\, n - 1$, of a total degree $\mathrm{Card}(M_N)\, n$, with coefficients that are rational functions in x, from $K = k(x)$, of degrees bounded by a constant depending only on N, and $l_i[f] : i = 1, \ldots, n$.

We claim now, that for an arbitrary $x_\alpha \in A_0$, arbitrary solutions $f_i = f_i(x)$ of $l_i[f_i] = 0$: $i = 1, \ldots, n$ and arbitrary polynomials $P_1, \ldots, P_n$ from $k[x]$, the *order* of $\mathfrak{D}_N(P_1, \ldots, P_n)$ at $x = x_\alpha$ is at least $\dim_k \mathfrak{M}_{N-1} \cdot \mathrm{ord}_{x = x_\alpha}(\sum_{i=1}^n P_i f_i) - C_2(N)$, where $C_2(N)$ depends only on F_γ.

To prove this we take an arbitrary basis $g_\delta : \delta \in M_{N-1}$ of $\mathfrak{M}_{N-1}$, $\mathrm{Card}(M_{N-1}) = \dim_k \mathfrak{M}_{N-1}$. Then, according to the graded structure, $\mathfrak{M}_1 \cdot \mathfrak{M}_{N-1} \subset \mathfrak{M}_N$, so for an arbitrary $g_\delta \in \mathfrak{M}_{N-1}$, $f_i \cdot g_\delta$ is an element of $\mathfrak{M}_N$. Thus $f_i \cdot g_\delta = \sum_{\gamma \in M_N} C_\delta^{i,\gamma} F_\gamma$ for constants $C_\delta^{i,\gamma} \in k$. Hence $\sum_{i=1}^n P_i f_i \cdot g_\delta = \sum_{i=1}^n P_i \{\sum_{\gamma \in M_N} C_\delta^{i,\gamma} F_\gamma\}$. Then we have $\mathrm{Card}(M_{N-1}) = \dim_k \mathfrak{M}_{N-1}$ vectors $\mathbf{C}_\delta = (C_\delta^{i,\gamma} : i = 1,\dots, n;\ \gamma \in M_N)$ that are linearly independent over k. If, on the contrary, $\mathbf{C}_\delta : \delta \in M_{N-1}$ are linearly dependent, then $\sum_{\delta \in M_{N-1}} d_\delta \mathbf{C}_\delta = 0$, for some constants $d_\delta \in k$, so that $\sum_{\delta \in M_{N-1}} d_\delta C_\delta^{i,\gamma} = 0$ for $i \leqslant n$, $\gamma \in M_N$ or $\sum_{\delta \in M_{N-1}} d_\delta f_i g_\delta = 0$ for all i. This implies, according to the linear independence of $g_\delta : \delta \in M_{N-1}$ over k, that all coefficients d_δ are zero.

We complete $\mathrm{Card}(M_{N-1})$ vectors $\mathbf{C}_\delta : \delta \in M_{N-1}$ to a nonsingular matrix A from $SL(\mathrm{Card}(M_N) \cdot n, k)$. We apply then the linear transformation induced by A to the Wronskian matrix in the denominator of the expression (3.3). Hence we reduce the determinant $W(\{P_i F_\gamma : i \leqslant n,\ \gamma \in M_N\})$ to the form, where the first $\mathrm{Card}(M_{N-1}) = \dim_k \mathfrak{M}_{N-1}$ columns, enumerated by $\delta \in M_{N-1}$, have the form $((d/dx)^{j-1} \cdot \{\sum_{i=1}^n \sum_{\gamma \in M_N} P_i \cdot F_\gamma \cdot C_\delta^{i,\gamma}\} : j = 1,\dots, \mathrm{Card}(M_N)\, n)^t$ or

$$\left(\left(\frac{d}{dx} \right)^{j-1} \left\{ \sum_{i=1}^{n} P_i f_i \cdot g_\delta \right\} : j = 1,\dots, \mathrm{Card}(M_N) \cdot n \right)^t$$

for $\delta \in M_{N-1}$.

From this representation of $W(\{P_i \cdot F_\gamma : i \leqslant n, \gamma \in M_N\})$ and (3.3) it follows that $\mathfrak{D}_N(p_1,\dots, p_n)$, considered as a function of the differential undetermined $p_1,\dots, p_n$, has the manifold

$$\{(p_1(x),\dots, p_n(x)) : p_1(x) f_1(x) + \cdots + p_n(x) f_n(x) \equiv 0\}$$

as a zero-manifold of $\mathfrak{D}_N(p_1,\dots, p_n)$ of the multiplicity $\mathrm{Card}(M_{N-1}) = \dim_k \mathfrak{M}_{N-1}$.

We can consider now the expansion of the differential polynomial $\mathfrak{D}_N(P_1,\dots, P_n)$ in powers of a new differential variable (the remainder function)

$$R = \sum_{i=1}^{n} P_i f_i.$$

Writing $P_n = -\sum_{i=1}^{n-1} P_i f_i / f_n + R/f_n$, and expanding $\mathfrak{D}_N(P_1,\dots, P_n)$ in powers of $(R/f_n)^{(m)} : m \geqslant 0$, we see that $\mathfrak{D}_N(P_1,\dots, P_n)$ is a sum of expressions of the form

$$(R/f_n)^{s_0} \cdot (\{R/f_n\}')^{s_1} \cdot \cdots \cdot (\{R/f_n\}^{(m)})^{s_m} \cdot \cdots \cdot \nabla \circ \mathfrak{D}_N(P_1,\dots, P_n),$$

where

$$\nabla = \prod_{m=0}^{L-1} (\partial / \partial P_n^{(m)})^{s_m}$$

for $L = \mathrm{Card}(M_N) \cdot n$ and always

$$s_0 + s_1 + \cdots + s_m + \cdots \geqslant \mathrm{Card}(M_{N-1}) = \dim_k \mathfrak{M}_{N-1}.$$

Let us assume now that $x_\alpha \in A_0$, that $f_i = f_i(x)$ are regular at $x = x_\alpha$ and $P_i = P_i(x)$ are polynomials from $k[x]$ (i.e., are regular at $x = x_\alpha$ as well): $i = 1,\ldots, n$. Then the expansion of $\mathfrak{D}_N(P_1,\ldots, P_n)$ in powers of derivatives of R implies

$$\begin{aligned} &\mathrm{ord}_{x=x_\alpha}(\mathfrak{D}_N(P_1,\ldots, P_n)) \\ &\qquad \geqslant \dim_k \mathfrak{M}_{N-1} \cdot \mathrm{ord}_{x=x_\alpha}\left(\sum_{i=1}^{n} P_i(x) f_i(x)\right) - C_2(N) \end{aligned} \tag{3.4}$$

for $C_2(N) > 0$ depending only on N and $l_1[\cdot],\ldots, l_n[\cdot]$, but independent of x_α and $f_1,\ldots,f_n$.

Remark 3.2. At this point of proof we remark that we do not assume $x = x_\alpha$ for $x_\alpha \in A_0$ to be a regular point of differential equations $l_1[y] = 0,\ldots, l_n[y] = 0$, but only a regular point of $f_1 = f_1(x),\ldots,f_n = f_n(x)$. If, however, $x = x_\alpha$ is at most a regular singularity of functions $g_\delta: \delta \in M_{N-1}$ and $F_\gamma: \gamma \in M_N$, then the bound (3.4) follows directly from the expansion of the determinant $W(\{P_i \cdot F_\gamma: i \leqslant n, \gamma \in M_N\})$.

To get an upper bound of $\mathrm{ord}_{x=x_\alpha}(\mathfrak{D}_N(P_1,\ldots, P_n))$ we notice that for polynomials $P_1(x),\ldots, P_n(x)$ from $k[x]$, $\mathfrak{D}_N(P_1,\ldots, P_n)$ is a rational function from $k(x)$ (with a bounded denominator). According to the description of $\mathfrak{D}_N(P_1,\ldots, P_n)$ above, we have the following upper bound of the degree of $\mathfrak{D}_N(P_1,\ldots, P_n)$ as a rational function from $k(x)$:

$$\deg(\mathfrak{D}_N(P_1,\ldots, P_n)) \leqslant \sum_{i=1}^{n} \deg(P_i) \dim_k \mathfrak{M}_N + C_3(N) \tag{3.5}$$

for $C_3(N) > 0$ depending only on $F_\gamma: \gamma \in M_N$.

According to the description of $\mathfrak{D}_N(P_1,\ldots, P_n)$, it is homogeneous polynomial in $P_i, P_i',\ldots, P_i^{(L-1)}$ of degree $\mathrm{Card}(M_N)$ for $L = \mathrm{Card}(M_N) \cdot n$ for every $i = 1,\ldots, n$; having coefficients from $k(x)$ of degrees bounded by $C_4(N)$. Consequently, for an arbitrary (finite) point a from k and $i = 1,\ldots, n$ we have

$$\mathrm{ord}_{x=a}(P_i) \cdot \dim_k \mathfrak{M}_N - C_5(N) \leqslant \mathrm{ord}_{x=a}(\mathfrak{D}_N(P_1,\ldots, P_n)) \tag{3.6}$$

for $C_5(N) = C_4(N) + \mathrm{Card}(M_N) \cdot n - 1$. Hence if $\mathfrak{D}_N(P_1,\ldots, P_n)$ is *nonzero*, the upper bound (3.5) and the lower bounds (3.4), (3.6) imply

$$\left\{ \sum_{\alpha \in A_0} \operatorname{ord}_{x=x_\alpha} \left(\sum_{i=1}^{n} P_i(x) f_i(x) \right) \right\} \cdot \dim_k \mathfrak{M}_{N-1} \tag{3.7}$$
$$+ \left\{ \sum_{i=1}^{n} \sum_{a \in A_i} \operatorname{ord}_{x=a}(P_i) \right\} \cdot \dim_k \mathfrak{M}_N$$
$$\leqslant \left\{ \sum_{i=1}^{n} \deg(P_i) \right\} \cdot \dim_k \mathfrak{M}_N + C_6(N).$$

Now, according to the Serre–Hilbert theorem, $\dim_k \mathfrak{M}_N / \dim_k \mathfrak{M}_{N-1} \to 1$ as $N \to \infty$. Hence, if $\mathfrak{D}_N(P_1, ..., P_n) \neq 0$ for $N \geqslant N_1(\varepsilon)$, then exponentiating (3.7) we get

$$\prod_{x_\alpha \in A_0} \left| \sum_{i=1}^{n} P_i(x) f_i(x) \right|_{x_\alpha} \cdot \prod_{i=1}^{n} \prod_{a \in A_i} |P_i|_a \tag{3.8}$$
$$\geqslant c_\varepsilon^0 \cdot \{|P_1|_\infty \cdot \cdots \cdot |P_n|_\infty\}^{-1-\varepsilon} = c_\varepsilon^0 \cdot \exp \left\{ (-1-\varepsilon) \cdot \left(\sum_{i=1}^{n} \deg(P_i) \right) \right\}$$

for $\sum_{i=1}^n \deg(P_i) \geqslant c_1(\varepsilon, l_1[\cdot], ..., l_n[\cdot])$.

It remains only to consider the case $\mathfrak{D}_N(P_1, ..., P_n) = 0$ for $N = N_1(\varepsilon)$. According to the representation (3.3) and the property of Wronskians, functions $P_i F_\gamma$ $(i \leqslant n,\ \gamma \in M_N)$ are linearly dependent over k: $\sum_{i=1}^n P_i(\sum_{\gamma \in M_N} c_{i,\gamma} F_\gamma) \equiv 0$ for constants $c_{i,\gamma} \in k$, not all zero. This implies, as we shall see, that polynomials $P_i(x)$ are related by a linear relation over $k(x)$ with coefficients having bounded degrees (depending only on F_γ). To prove this we choose the basis $\{h_\zeta\}_{\zeta \in L_N}$ of $\mathfrak{M}_N$ over $K = k(x)$, so that all elements of $\mathfrak{M}_N$ are expressed in this basis with coefficients from $k(x)$ of degrees bounded by some m, $m = m(N)$. Then $\sum_{\gamma \in M_n} c_{i,\gamma} F_\gamma = \sum_{\zeta \in L_N} e_i^\zeta h_\zeta$ and $\sum_{i=1}^n P_i(\sum_{\zeta \in L_N} e_i^\zeta h_\zeta) \equiv 0$ for $e_i^\zeta \in K$ of degrees $\leqslant m$, and not all e_i^ζ $(\zeta \in L_N$, $i = 1, ..., n)$ are identically zero. Since $\{h_\zeta\}_{\zeta \in L_N}$ are linearly independent over K, we arrive to at least one nontrivial linear relations among P_i: $\sum_{i=1}^n e_i P_i = 0$ for $e_i \in K$ of degrees $\leqslant m$ (where $e_i = e_i^\zeta$ for some $\zeta \in L_N$), $i = 1, ..., n$. We can assume, without loss of generality, that $e_n \neq 0$ and obtain $A(x) P_n(x) = \sum_{i=1}^n B_i(x) P_i(x)$ for polynomials $A(x) \not\equiv 0$, $B_i(x)$ from $k[x]$ of degrees $\leqslant m' = m'(N)$: $i = 1, ..., n$. Then

$$\sum_{i=1}^{n} P_i(x) f_i(x) = A(x)^{-1} \cdot \left\{ \sum_{i=1}^{n-1} P_i(x) \cdot (A(x) f_i(x) + B_i(x) f_n(x)) \right\}.$$

If we denote $h_i(x) = A(x) f_i(x) + B_i(x) f_n(x)$: $i = 1, ..., n-1$, then we obtain for an arbitrary $x_\alpha \in A_0$

$$\operatorname{ord}_{x=x_\alpha} \left(\sum_{i=1}^{n} P_i(x) f_i(x) \right) \leqslant \operatorname{ord}_{x=x_\alpha} \left(\sum_{i=1}^{n-1} P_i(x) h_i(x) \right) - m'(N). \tag{3.9}$$

Since the degrees of $A(x)$, $B_i(x)$ are bounded by m', we have $h_i(x) = \sum_{j=0}^{m'}\{b_j x^j f_i(x) + d_j x^j f_n(x)\}$ for some constants $b_j, d_j: j = 0,..., m'$ from k. This shows that $h_i(x)$ is a solution of a linear differential equation $L_i[h_i] = 0$ over K of the order $\leqslant m'(\mathrm{ord}(l_i) + \mathrm{ord}(l_n))$, uniquely and effectively determined by $l_i[\cdot]$ and $l_n[\cdot]: i = 1,..., n-1$ and $m' = m'(N)$ only. Thus, according to (3.9), if $\mathfrak{D}_N(P_1,..., P_n) = 0$ (for $N = N_1(\varepsilon)$), then the proof of (3.1), (3.2) for a given n is reduced to the case of $n-1$ functions. The functional approximation theorem is proved by induction.

The proof of the functional "Schmidt's theorem" above gives at the same time a theorem on simultaneous approximations that we formulated in terms of the norming $|\cdot|_\infty$.

Simultaneous Approximation Theorem. *For every $\varepsilon > 0$ and linear differential equations $l_1[y] = 0,..., l_n[y] = 0$ over $k(x)$ there exist constants $\bar{c}_\varepsilon^0$ and $\bar{c}_\varepsilon^1$ such that the following is true. For arbitrary solutions $f_1 = f_1(x),..., f_n = f_n(x)$ of equations $l_1[f_1] = 0,..., l_n[f_n] = 0$, respectively, that are regular at $x = \infty$, we have*

$$\prod_{i=1}^{n} \|Q(x) f_i(x)\|_\infty > \bar{c}_\varepsilon^0 \cdot |Q|_\infty^{-1-\varepsilon}$$

for arbitrary polynomial $Q(x)$ from $k[x]$ and $\|Q \cdot f_i\|_\infty = \min_{P(x)\in k[x]} |Q \cdot f_i - P|_\infty$, provided that $|Q|_\infty > \bar{c}_\varepsilon^1$.

Here $\bar{c}_\varepsilon^0$ and $\bar{c}_\varepsilon^1$ depend effectively only on $l_1[\cdot],..., l_n[\cdot]$ and ε.

The simultaneous approximation theorem can be also directly deduced from the functional approximation theorem using the Khintchine transfer theorem in the functional case.

We conjecture that $\varepsilon = 0$ in the statement of the functional approximation theorem. This is equivalent to the "almost perfectness" of arbitrary multipoint Hermite–Padé approximations to functions $f_1(x),..., f_n(x)$.

Remark. The conjecture that $\varepsilon = 0$ is proved by the authors in the following cases: (i) where $f_i(x)$ satisfy linear differential equations with constant coefficients; (ii) when $f_i(x)$ satisfy hypergeometric equations, e.g., Bessel equations; (iii) when $f_i(x)$ are algebraic functions from the cyclic extension $\mathbb{C}(\sqrt[m]{x-x_0})$ for some m. In Section 5 we prove that $\varepsilon = 0$ in another case, when $f_i(x)$ belongs to the algebraic extension $\mathbb{K}$ of $\mathbb{C}(x)$ such that $[\mathbb{K}:\mathbb{C}(x)] = d$ and $n = d-1$ and $\sum_{i=1}^{n} \deg(P_i)$ is replaced by $n \cdot \max\{\deg(P_i): i = 1,..., n\}$.

Even though we cannot prove now that $\varepsilon = 0$ in the general case, the proof of the functional approximation theorem shows that always ε can be replaced by a function tending to zero as $\sum_{i=1}^{n} \deg(P_i) \to \infty$,

$$\sum_{x_\alpha \in A_0} \mathrm{ord}_{x=x_\alpha}\left(\sum_{i=1}^{n} P_i(x) f_i(x)\right) \leqslant D + C^0 \cdot D^{(m-1)/m}$$

for

$$D=\sum_{i=1}^{n}\deg(P_i)\geqslant C^1 \qquad \text{and} \qquad m=\operatorname{ord}(l_1)\cdots\operatorname{ord}(l_n).$$

Moreover, the function ε and the exponent $(m-1)/m$ has homological sense. In general m is determined using the Hilbert polynomial $P(N)$ of the graded ring $\sum_{N\geqslant 0}\mathfrak{M}_N$, where

$$\dim_k \mathfrak{M}_N = P(N) \qquad \text{for} \quad N\geqslant N_0$$

and the polynomial $P(N)$ is of degree $m-1$. Hence one has $m\leqslant \operatorname{ord}(l_1)\cdot\cdots\cdot\operatorname{ord}(l_n)$ in the general case.

For special linear differential operators $l_1[\cdot],\dots,l_n[\cdot]$ the exponent m is considerably smaller than $\operatorname{ord}(l_1)\cdot\cdots\cdot\operatorname{ord}(l_n)$, and is expressed in terms of the dimension of the Picard–Vessiot group G of M. The case of algebraic functions is particularly interesting because in this case always $m=2$. Hence we obtain

COROLLARY 3.1. *For arbitrary algebraic functions $f_1(x),\dots,f_n(x)$ and polynomials $P_1(x),\dots,P_n(x)$ from $k[x]$ we have*

$$\sum_{x_\alpha\in A_0}\operatorname{ord}_{x=x_\alpha}(P_1(x)f_1(x)+\cdots+P_n(x)f_n(x))$$
$$\leqslant\sum_{i=1}^{n}\deg(P_i)+C_7\sqrt{\sum_{i=1}^{n}\deg(P_i)+1}$$

if $P_1(x)f_1(x)+\cdots+P_n(x)f_n(x)\not\equiv 0$. Here $P_1(x)f_1(x)+\cdots+P_n(x)f_n(x)$ is assumed to be regular at $x=x_\alpha$ for all $x_\alpha\in A_0$, the constant C_7 is effectively computable in terms of the function field $M=k(x,f_1(x),\dots,f_n(x))$ and $\operatorname{Card}(A_0)$. *Moreover, C_7 depends linearly on* $\operatorname{Card}(A_0)$.

For nonlinear algebraic differential equations the Roth and Schmidt theorems can also be proved.

4. The Effective Roth Theorem

In this chapter we specialize the arguments of Section 3 to the case of $n=2$, i.e., to approximations of solutions of linear differential equations by rational functions. In this way we arrive to the "effective Roth" theorem (see [36] for the "noneffective Roth" theorem for algebraic functions). Such specialization is necessary to reveal the importance of the Ricatti equations of Section 1 and the crucial role of the concept of differential denomination of a differential polynomial introduced by Kolchin in [18] and used by Osgood in [25–28]. We recall the definition of differential denomination from [18].

DEFINITION 4.1. A differential polynomial $P(y) = P(y, y', \ldots, y^{(m-1)})$ has a differential denomination $\leqslant d$, if for two differential indeterminates u and v, $P(u/v) \cdot v^d$ is a differential polynomial in u and v.

We also discuss the structure of the graded ring $\sum_{N \geqslant 0} \mathfrak{M}_N$ from Section 2 from the point of view of the reducibility of linear differential equations; for a few of the existing general results in this direction see [19].

Let $f(x)$ be a solution of a linear differential equation of dth order with rational function coefficients. Let us denote by $L[u]$ the linear differential operator of dth order, which annihilates $f(x)$,

$$L[u] = u^{(d)} + a_{d-1} u^{(d-1)} + \cdots + a_0 u, \tag{4.1}$$

where $a_i(x) \in k(x)$ for the algebraically closed field k of characteristic $0, i = 0, \ldots, d-1$. Let us denote by $f_1(x), \ldots, f_d(x)$ the fundamental system of solutions of an dth order linear differential equation

$$L[f(x)] = 0.$$

There is a famous representation of the operator $L[u]$ as a ratio of two Wronskians,

$$L[u] = \frac{W(u, f_1, \ldots, f_d)}{W(f_1, \ldots, f_d)}. \tag{4.2}$$

As in Section 2, for every $N \geqslant 0$ we consider a vector space $\mathfrak{M}_N$ over k generated by all monomials $f_1^{m_1} \cdot \cdots \cdot f_d^{m_d}$ in $f_1, \ldots, f_d$ of degree N: $m_1 + \cdots + m_d = N$. According to the Picard–Vessiot theory from Section 2, the vector space $\mathfrak{M}_N$ coincides with the space of solutions of a linear differential equation $L^{(N)}[y] = 0$ over $k(x)$, depending only on $L[y]$ and N. According to the Serre–Hilbert theorem (see, e.g., [2]) $\dim_k \mathfrak{M}_N = P_{m-1}(N)$ for an integer valued polynomial $P_{m-1}(N)$ of degree $m-1$ in N, and $N \geqslant N_0$. We choose a basis $f_\mu : \mu \in M_N$ of $\mathfrak{M}_N$ over k, $\mathrm{Card}(M_N) = \dim_k \mathfrak{M}_N$.

In the case, when $m < d$, or $\mathrm{Card}(M_N)$ ($= \dim_k \mathfrak{M}_N$) $< \binom{N+d-1}{d-1}$, there are nontrivial homogeneous algebraic relations between $f_1, \ldots, f_d$. This case can be characterized as "degenerate" or "reducible." The degeneracy and the precise determination of the dimension m in this case can be achieved only in terms of the Galois group G of the corresponding Picard–Vessiot extension $M = K\langle f_1, \ldots, f_d \rangle$ of $K = k(x)$, see [19]. Classical literature on differential equations (Fuchs, Fano,..., etc.) seem to support a conjecture that in a degenerate case, solutions of $L[f]=0$ can be represented as products of solutions of linear differential equations of smaller order and (or) algebraic and exponential functions. This conjecture was proved in classical literature only for $m \leqslant 2$. We always have $m \leqslant d$ and if $d=2$, then always $\mathrm{Card}(M_N)=\dim_k \mathfrak{M}_N = N+1$. If $d=3$, then the degeneracy of $L[f]=0$ implies that either solutions of the equations $L[f]=0$ are products of solutions of the second order linear differential equations (when there are nontrivial quadratic relations between solutions of $L[f]=0$), or else all solutions of $L[f]=0$ are algebraic functions.

As in Section 3 we introduce now a nonlinear differential operator $\mathfrak{D}_N[u]$, which is a differential polynomial in $u, u',\dots, u^{(L-1)}$, $L=\dim_k \mathfrak{M}_{N+1} + \dim_k \mathfrak{M}_N$ with coefficients that are rational in x, and such that $\mathfrak{D}_N[u]$ has a denomination in u at most $\dim_k \mathfrak{M}_N + \dim_k \mathfrak{M}_{N+1}$ and such that the function f is a singular solution of $\mathfrak{D}_N[u]$ of order at least $\dim_k \mathfrak{M}_N$. The explicit construction of this differential operator $\mathfrak{D}_N[u]$ enables us to prove the Roth's theorem for $f(x)$ following the Kolchin's approach [18].

THEOREM 4.2. *For $N \geqslant 0$ we introduce the following differential operator $\mathfrak{D}_N[u]$,*

$$\mathfrak{D}_N[u] = \frac{W(\{uf_\mu : \mu \in M_N\} \cup \{f_\eta : \eta \in M_{N+1}\})}{W(\{f_\mu : \mu \in M_N\}) \cdot W(\{f_\eta : \eta \in M_{N+1}\})}. \tag{4.3}$$

Then $\mathfrak{D}_N[u]$ is a differential polynomial in $u, u',\dots, u^{(L-1)}$ for $L=\mathrm{Card}(M_{N+1})+\mathrm{Card}(M_N)$ such that the following properties are satisfied:

(i) *$\mathfrak{D}_N[u]$ is a differential polynomial in $u, u',\dots, u^{(L-1)}$ with coefficients being rational functions in x. These rational functions are polynomials over $\mathbb{Z}$ in coefficients a_i of the linear operator $L[u]$ and their derivatives $a_i^{(j)}$. The orders j of $a_i^{(j)}$ and degrees in $a_i^{(j)}$ are bounded by constants depending only on d and N.*

(ii) *Every solution $f(x)$ of the linear differential equation $L[f(x)]=0$ is a singular solution of $\mathfrak{D}_N[u]=0$ of order at least* $\mathrm{Card}(M_N)$. *This means that for every differential monomial $\nabla = (\partial/\partial u)^{a_0}(\partial/\partial u')^{a_1}\cdots(\partial/\partial u^{(\alpha)})^{a_\alpha}$, $\alpha \leqslant L-1$ of order $a_0+a_1+\cdots+a_\alpha \leqslant$* $\mathrm{Card}(M_N)-1=\dim_k \mathfrak{M}_N - 1$ *the function $u=f(x)$ is a solution of $\nabla \circ \mathfrak{D}_N[u]=0$.*

(iii) *The denomination of the differential polynomial $\mathfrak{D}_N[u]$ in u is at most $L=$* $\mathrm{Card}(M_N)+\mathrm{Card}(M_{N+1})$.

Proof. Part (i) is a simple corollary of the properties of Wronskian; the rationality of the coefficients of $\mathfrak{D}_N[u]$ is a consequence of the invariance of its coefficients under an arbitrary linear transformation of the fundamental system $f_1(x),\dots,f_d(x)$ of solutions of $L[f(x)]=0$.

The proof of (ii) follows from an obvious observation that for every $\mu \in M_N$, and any solution u of $L[u]=0$, $u \cdot f_\mu$ is a linear combination of functions $f_\eta : \eta \in M_{N+1}$ generating $\mathfrak{M}_{N+1}$ (since $u=\sum_{i=1}^n c_i f_i$ and so $u \cdot \mathfrak{M}_N \subseteq \mathfrak{M}_{N+1}$). Then $\nabla \circ W(\{uf_\mu\} \cup \{f_\eta\})$ is represented as a sum of determinants each of which has at least one column as $(uf_{\mu_0})^{(j-1)} : j=1,..., L$ and $\mathrm{Card}(M_{N+1})$ columns $(f_\eta)^{(j-1)} : j=1,..., L$, for all $\eta \in M_{N+1}$.

For the proof of part (iii) we put $u=Y/Z$ and use the well-known property of Wronskians: $W(h \cdot h_1,..., h \cdot h_m)=h^m \cdot W(h_1,..., h_m)$. Then we have

$$\mathfrak{D}_N \left| \frac{Y}{Z} \right| = Z^{-\mathrm{Card}(M_N)-\mathrm{Card}(M_{N+1})} \cdot \frac{W(\{Y \cdot f_\mu\} \cup \{Z \cdot f_\eta\})}{W(\{f_\mu\}) \cdot W(\{f_\eta\})}$$

so that $\mathfrak{D}_N[u]$ has a denomination in u at most $\mathrm{Card}(M_N)+\mathrm{Card}(M_{N+1})$.

The proof of the Main Theorem is completed.

In order to show that $\mathfrak{D}_N[u]$ is not annihilating "too many" rational functions we need the following almost trivial

LEMMA 4.3. *Let $h_1(x),..., h_n(x)$ and $g_1(x),..., g_m(x)$ be arbitrary families of formal power series. Then among the ratios*

$$\frac{\sum_{i=1}^n c_i h_i(x)}{\sum_{j=1}^m d_j g_j(x)} \tag{4.4}$$

for arbitrary constants $c_1,..., c_n$ and $d_1,..., d_m$, there can appear rational functions of bounded degrees only. Moreover, degrees of rational functions among expressions (4.4) *are bounded by a constant depending only on $h_1,..., h_n$ and $g_1,..., g_m$.*

For the proof of Lemma 4.3 we choose among the functions $h_1,..., h_n$, $g_1,..., g_m$ the maximal set of functions linearly independent over $k(x)$. If this basis is, say $\varphi_1(x),..., \varphi_p(x)$, then every linear combination $\sum_{i=1}^n c_i h_i(x)$ and $\sum_{j=1}^m d_j g_j(x)$ can be represented as a linear combination $\sum_{l=1}^p \alpha_l(x)\, \varphi_l(x)$ or $\sum_{l=1}^p \beta_l(x)\, \varphi_l(x)$, resectively, with rational functions $\alpha_l(x)$, $\beta_l(x)$ having bounded degrees (depending only on $h_1,..., h_n$, $g_1,..., g_m$). Let now $r(x)$ be a rational function and $r(x)=\sum_{i=1}^n c_i h_i(x)/\sum_{j=1}^m d_j g_j(x)$, so that $\sum_{l=1}^p \alpha_l(x)\, \varphi_l(x)=r(x) \cdot (\sum_{l=1}^p \beta_l(x)\, \varphi_l(x))$ and $\alpha_l(x)-r(x)\, \beta_l(x)=0$ for all $l=1,..., p$. Here not all $\alpha_l(x)$, $\beta_l(x)$ are zero, and hence the degrees of rational functions represented in the form (4.4) are bounded.

COROLLARY 4.4. *If a rational function $u=P/Q$ (where P and Q are relatively prime polynomials) satisfies the differential equation $\mathfrak{D}_N[u]=0$, then the degrees of P and Q are bounded by a constant depending only on N and linear differential operator* (4.1).

Proof. According to the main property of Wronskians $\mathfrak{D}_N[u]=0$ if and only if functions $uf_\mu, f_\eta: \mu\in M_N$, $\eta\in M_{N+1}$ are linearly dependent over k. This implies that $\sum_{\mu\in M_N} c_\mu\cdot uf_\mu-\sum_{\eta\in M_{N+1}} d_\eta f_\eta=0$ for constants c_μ, d_η, where not all of c_μ are zero. Hence

$$u=\frac{\sum_{\eta\in M_{N+1}} d_\eta f_\eta}{\sum_{\mu\in M_N} c_\mu f_\mu}.$$

Then Lemma 4.3 bounds the degrees of rational functions.

Theorem 4.2 and Corollary 4.4 imply

THEOREM 4.5 (Effective Roth theorem). *Let $f(x)$ be a solution of a linear differential equation with rational function coefficients. Then for any $\varepsilon>0$ there exists an effective constant $C_\varepsilon>0$ such that*

$$\sum_{x_0\in A} \operatorname{ord}_{x=x_0}\left(f(x)-\frac{P(x)}{Q(x)}\right)<\deg(P)+\deg(Q)(1+\varepsilon)$$
$$+\operatorname{Card}(A)\cdot C_\varepsilon,$$

if $\deg(P)+\deg(Q)>C_\varepsilon$, *and polynomials $P(x)$ and $Q(x)$ are relatively prime.*

Here A is a set of (finite) points $x=x_0$, where $f(x)$ is regular, and the constant C_ε can be effectively determined in terms of ε and the linear differential equation satisfied by $f(x)$.

Proof. Let $u=P/Q$ be an approximation to $f(x)$ for relatively prime polynomials $P(x)$ and $Q(x)$. Then according to (ii) of the Theorem 4.2 and the expression (4.3) of $\mathfrak{D}_N[u]$ we obtain

$$\operatorname{ord}_{x=x_0}(\mathfrak{D}_N[u])\geqslant \operatorname{Card}(M_N)\cdot\operatorname{ord}_{x=x_0}(u-f)-\gamma_1(N),$$

where $\gamma_1(N)$ depends on $\operatorname{Card}(M_N)$, $\operatorname{Card}(M_{N+1})$.

On the other hand, according to Corollary 4.4, if $\deg(P)+\deg(Q)>\gamma_2(N)$, then $\mathfrak{D}_N[P/Q]\neq 0$. Hence $\mathfrak{D}_N[u]$ is a nonzero differential polynomial in P/Q of denomination of at most $\operatorname{Card}(M_N)+\operatorname{Card}(M_{N+1})$ with coefficients being rational functions in x with degrees depending only on N. Thus $\mathfrak{D}_N[u]=Q^{-L}\cdot D_N[x;P,Q]$ for $L=\operatorname{Card}(M_N)+\operatorname{Card}(M_{N+1})$. Here $D_N[x;P,Q]$ is a differential polynomial in P and Q, of the degree $\operatorname{Card}(M_N)$ in P, and of the degree $\operatorname{Card}(M_{N+1})$ in Q, with coefficients that are rational functions in x of degrees at most $\gamma_3(N)$. Thus

$$\sum_{x_0\in A}\operatorname{ord}_{x=x_0}(\mathfrak{D}_N[u])\leqslant\deg(P)\cdot\operatorname{Card}(M_N)$$
$$+\deg(Q)\cdot\operatorname{Card}(M_{N+1})+\gamma_3(N).$$

This gives that

$$\deg(P)\cdot \operatorname{Card}(M_N)+\deg(Q)\cdot \operatorname{Card}(M_{N+1})$$
$$\geqslant \sum_{x_0\in A}\{\operatorname{ord}_{x=x_0}(u-f)\cdot \operatorname{Card}(M_N)-\gamma_4(N)\},$$

then

$$\sum_{x_0\in A}\operatorname{ord}_{x=x_0}(f-P/Q)\leqslant \deg(P)+\deg(Q)\cdot\frac{\operatorname{Card}(M_{N+1})}{\operatorname{Card}(M_N)}$$
$$+\operatorname{Card}(A)\cdot\gamma_4(N)/\operatorname{Card}(M_N).$$

Since $\operatorname{Card}(M_N)$ is a polynomial in N, the Theorem 4.5 is proved.

Let us present the expression for a few first operators $\mathfrak{D}_N[u]$. First of all according to (4.2), $\mathfrak{D}_0[u]=L[u]$ for a linear differential operator $L[u]$ in (4.1). Now let $d=2$ and $L[u]=u''+bu'+cu$. Then

$$D_1[u]=4\left(\frac{d}{dx}L[u]\right)^2-3\cdot L[u]\cdot\frac{d^2}{dx^2}L[u]$$
$$+b\cdot L[u]\cdot\frac{d}{dx}L[u]+(9c-2b^2-6b')\cdot L[u]^2.$$

In general, for $d=2$, $\mathfrak{D}_N[u]$ is a homogeneous function of degree $N+1$ in $L[u]$ and its derivatives up to the order $2N$, with coefficients that are polynomials in b, c and their derivatives.

The explicit computation of constants $\gamma_1(N),\ldots,\gamma_4(N)$ in the proof of Theorem 4.5 above is not difficult. For example, when $d=2$, so that $f(x)$ satisfies a linear differential equation of the second order, then $\operatorname{ord}_{x=x_0}(f(x)-P(x)/Q(x))\leqslant \deg(P)+\deg(Q)+c_1\sqrt{\deg(P)+\deg(Q)}+c_2$ with $c_1>0$, $c_2>0$ effectively determined in terms of the differential equation (4.1) satisfied by $f(x)$. Here the upper bound on c_1 is $c_1\leqslant D-1$, where D is the maximal degree of the coefficients of (4.1).

We conclude this chapter with the specialization of the $\varepsilon=0$ conjecture from Section 3. In the case of rational approximations to a single function, this conjecture reads as follows. The continued fraction expansion of the solution of a (linear) differential equation with rational function coefficients, has bounded (degrees of) elements, see Section 3 for partial results in this direction.

5. Identities between Solutions of the Ricatti Equations; Bäcklund Transformations and Their Applications

Wronskian formalism and the introduction of the auxiliary Ricatti equation, that lead to the proof of the theorems of Schmidt and Roth, can be used further. They yield a proof of the conjecture $\varepsilon = 0$ from Section 3 (the boundedness of the degrees of the elements of the continued fraction expansion) for particular classes of functions; especially for algebraic functions. One of the methods we propose, is based on the identities satisfied by Wronskians. An analogy with the invariant theory suggests the consideration of bideterminants similar to Wronskians and the study of nonlinear identities for bideterminants analogous to the laws of compositions of Bäcklund transformations (BTs) considered as Wronskians. We present a few examples of this approach, concentrating on the algebraic functions $f(x)$.

To explain our motivation we present a brief description of the universal identities satisfied by BTs, following our papers [7–9]. Using the notations introduced in Section 1, we study the action of BTs on the pseudopotential σ. Let σ_{λ_i} denote an elementary BT of σ corresponding to the addition of an apparent singularity at $\lambda = \lambda_i$ (i.e., $\sigma_{\lambda_i} = \psi(\lambda_i) \cdot \sigma$ where $\psi(\lambda_i)$ is an eigenfunction of (1.1) corresponding to the pseudopotential σ). The main two universal identities satisfied by iterations of elementary BTs, as presented in [8, 9], are the following:

$$(\lambda_2 - \lambda_3)\,\sigma_{\lambda_1} \cdot \sigma_{\lambda_2,\lambda_3} + (\lambda_3 - \lambda_1)\,\sigma_{\lambda_2} \cdot \sigma_{\lambda_3,\lambda_1} + (\lambda_1 - \lambda_2)\,\sigma_{\lambda_3} \cdot \sigma_{\lambda_1,\lambda_2} = 0; \tag{5.1}$$

and

$$\begin{aligned}(\lambda_1 - \lambda_4)(\lambda_2 - \lambda_3)\,\sigma_{\lambda_1,\lambda_4} \cdot \sigma_{\lambda_2,\lambda_3} + (\lambda_3 - \lambda_1)(\lambda_2 - \lambda_4)\,\sigma_{\lambda_1,\lambda_3} \\ \times \sigma_{\lambda_2,\lambda_4} + (\lambda_2 - \lambda_1)(\lambda_4 - \lambda_3)\,\sigma_{\lambda_1,\lambda_2} \cdot \sigma_{\lambda_4,\lambda_3} = 0.\end{aligned} \tag{5.2}$$

These laws of composition of Bäcklund transformations have the form of the laws of addition on (infinite dimensional) Jacobian varieties, if the action of the BT: $B_\lambda : \sigma \mapsto \sigma_\lambda$ is represented as a translation operator using the action of the vertex operator $X(\lambda)$ from [17, 23], that corresponds to the Lie algebra $A_\infty^{(1)}$. Then $\sigma_\lambda = \exp\{\sum_{k=1}^{\infty} \lambda^{-k}(\partial/\partial x_k)\}\,\sigma$ acts on the space of functions in the auxiliary variables x_k: $k = 1, 2,\ldots$. The variables x_k have an obvious interpretation as dynamic variables in the linear spectral problem (1.1), and a more nontrivial interpretation as Newton symmetric functions $x_k = \sum_i \alpha_i^{-k}$: $k = 1, 2,\ldots,$ in parameters α_i. Then S-functions are pseudopotentials, and the BT B_λ replaces the set $\{\alpha_i\}$ with $\{\alpha_i\} \cup \{\lambda\}$ (i.e., if $\sigma = \sigma(\sum \alpha_i^{-1}, \sum \alpha_i^{-2},\ldots)$, then $\sigma_\lambda = \sigma(\sum \alpha_i^{-1} + \lambda^{-1}, \sum \alpha_i^{-2} + \lambda^{-2},\ldots)$). In this interpretation of σ_λ, the laws of addition (5.1) and (5.2) in the case of Jacobians of finite genus assume the familiar forms of the law of addition for

Riemann θ-functions (cf. (14, 15]). Though identities (5.1) and (5.2) have such a deep interpretation, their derivation is very simple. According to the definition of BTs in terms of Wronskians, given above, the laws of composition (5.1) and (5.2) are equivalent to the following identities:

$$aW(b, c) + bW(c, a) + cW(a, b) = 0, \qquad (5.1')$$

$$W(a, d)\, W(c, b) + W(a, c)\, W(b, d) + W(a, b)\, W(d, c) = 0, \qquad (5.2')$$

where a, b, c, d are arbitrary functions. Identities (5.1′) and (5.2′) are the simplest in the infinite sequence of identities satisfied by alternant determinants. These identities can be used to construct nonlinear differential operators arising from the antisymmetric part of the tensor powers of the module of solutions of a linear differential equation. While the Ricatti equation that corresponds to a symmetric tensor power of this module gives us the Roth theorem, the antisymmetric power can furnish much more. We present an example of a particular case of Schmidt's theorem for algebraic functions, with the best possible exponent. For this we start with an algebraic function $\zeta = \zeta(x)$ of degree $d \geqslant 2$. Let $\zeta^{(1)} = \zeta, \zeta^{(2)},..., \zeta^{(d)}$ be functions algebraically conjugate to ζ: they are distinct solutions of an irreducible algebraic equation over $k(x)$ of degree d satisfied by ζ. We take an arbitrary basis $f_1 = f_1(x),..., f_d = f_d(x)$ of the function field $k(x, \zeta)$ over $k(x)$ and we put $n = d - 1$. We introduce analogs of Wronskians adjusted to this particular problem. For every $i = 1,..., d$ we denote by $f_1^{(i)},..., f_d^{(i)}$ the functions conjugated to $f_1,..., f_d$ under the isomorphisms of the field $k(x, \zeta): \zeta \to \zeta^{(i)}$: $i = 1,..., d$. Let us put for $i = 1,..., d$ we have $\Delta_i = \det(f_j^{(k)} : j = 1,..., n;$ $k = 1,..., d, k \neq i)$. Then we introduce the following auxiliary differential polynomial in the differential undetermined $p_1,..., p_n$,

$$\mathfrak{D}(p_1,..., p_n) = W(\Delta_i \cdot (p_1 f_1^{(i)} + \cdots + p_n f_n^{(i)}): \\ i = 1,..., n)/\det(f_j^{(k)} : j = 1,..., d;\, k = 1,..., d)^n.$$

First, we have to prove that $\mathfrak{D}(p_1,..., p_n)$ is a differential polynomial in $p_1,..., p_n$ with coefficients rational in x. For this we must verify that $\mathfrak{D}(p_1,..., p_n)$ is invariant under the action of the monodromy group of the equation satisfied by ζ. Since ζ is an algebraic function of degree d, this group is a subgroup of the permutation group S_d, so it is sufficient to verify that $\mathfrak{D}(p_1,..., p_n)$ is invariant under permutations of $(1,..., d)$.

To see the effect of the permutations of $(1,..., d)$ on $\mathfrak{D}(p_1,..., p_n)$ we use the identities satisfied by Δ_i, which are direct analogs of the laws of addition (5.1′) and (5.2′), in the case when derivative is replaced by the conjugation. We obtain

$$\sum_{i=1}^{d} f_k^{(i)} \Delta_i (-1)^i = 0 \qquad (5.3)$$

for all $k = 1, \dots, n$ $(=d-1)$. As above, this identity follows from the remark that the matrix

$$\begin{pmatrix} f_1^{(1)} f_2^{(1)} & \cdots & f_n^{(1)} x^1 \\ f_1^{(2)} f_2^{(2)} & \cdots & f_n^{(2)} x^2 \\ \vdots & \ddots & \vdots \\ f_1^{(d)} f_2^{(d)} & \cdots & f_n^{(d)} x^d \end{pmatrix}$$

has rank $\leqslant n$, if $x^i = f_k^{(i)} : i = 1, \dots, d$, for every $k = 1, \dots, n$.

For the permutation π_{ij} of $(1, \dots, d)$ we have $\pi_{ij}(\Delta_k) = -\Delta_k$ if $i \neq k$, $j \neq k$ and $\pi_{ij}(\Delta_i) = (-1)^{i+j+1} \Delta_j$. Let $R^{(i)} = p_1 f_1^{(i)} + \cdots + p_d f_d^{(i)}$. Then for $i = 1, \dots, n$ we have $\pi_{id}(W(\{\Delta_i R^{(i)} : i = 1, \dots, n\})) = (-1)^{i+1} \times W(\Delta_1 R^{(1)}, \dots, \Delta_{i-1} R^{(i-1)}, \Delta_d R^{(d)}, \Delta_{i+1} R^{(i+1)}, \dots, \Delta_n R^{(n)})$. However, $\Delta_d R^{(d)} = -\sum_{i=1}^n \Delta_i R^{(i)} (-1)^{d+i}$ according to (5.3), so that $\pi_{id}(W(\{\Delta_i R^{(i)}\})) = (-1)^d \cdot W(\{\Delta_i R^{(i)}\})$. Next, for $i < j \leqslant n$, $\pi_{ij}(W(\{\Delta_i R^{(i)}\})) = (-1)^d \cdot \pi_{ij}(W(\{\Delta_i R^{(i)}\}))$. Hence $\mathfrak{D}(p_1, \dots, p_n)$ is a differential polynomial in $p_1, \dots, p_n$ of total degree n and of order $\leqslant n-1$. Moreover, according to the construction of $\mathfrak{D}(p_1, \dots, p_n)$, this polynomial is not identically zero. This immediately implies that $\varepsilon = 0$ in one particular case of the Schmidt theorem. For the detailed proof see the proof of the theorem of Section 3 with $\mathfrak{D}_N(P_1, \dots, P_n)$ replaced by $\mathfrak{D}(p_1, \dots, p_n)$.

THEOREM 5.1. *Let $\zeta = \zeta(x)$ be an algebraic function of degree $d \geqslant 3$. Then for arbitrary algebraic functions $f_1(x), \dots, f_n(x)$; $n = d-1$, from $k(x, \zeta)$ and arbitrary (finite) x_0, where $\zeta(x)$ is regular, we have the following bound:*

$$\begin{aligned} &\operatorname{ord}_{x=x_0}(p_1(x) f_1(x) + \cdots + p_n(x) f_n(x)) \\ &\qquad \leqslant n \cdot \max\{\deg(p_i) : i = 1, \dots, n\} + C_1 \end{aligned}$$

for arbitrary polynomials $p_1(x), \dots, p_n(x)$ and a constant $C_1 > 0$ depending only on $\zeta(x)$ (and not on x_0).

For $n = 3$ this result was explicitly proved by Osgood in [25] using the Ricatti equation satisfied by a cubic algebraic function.

REFERENCES

1. R. J. ARMS AND A. EDREI, "The Padé Tables and Continued fractions Generated by Totally Positive Sequences," Mathematical Essays dedicated to A. J. Macintyre, pp. 1–21, Ohio Univ. Press, Athens, 1970.
2. M. F. ATIYAH AND I. G. MACDONALD, "Introduction to Commutative Algebra," Addison–Wesley, Reading, Mass., 1969.

3. R. M. MIURA (Ed.), Bäcklund transformations, *in* "The Inverse Scattering Method, Solitons and Their Applications," Lecture Notes in Mathematics, Vol. 515, Springer-Verlag, Berlin/New York, 1976.
4. A. BAKER, "Transcendental Number Theory," Cambridge Univ. Press, London, 1979.
5. F. CALOGERO AND A. DEGASPERIS, "Spectral Transform and Solitons," North-Holland Studies in Mathematics and Its Applications, Vol. 13, North-Holland, Amsterdam, 1982.
6. E. B. CHRISTOFFEL, Uber die Gaußishe Quadratur und eine Verallgemeinerung derselben, *J. Reine Angew. Math.* **55** (1858), 61–82.
7. D. V. CHUDNOVSKY AND G. V. CHUDNOVSKY, Multisoliton formula for completely integrable two-dimensional systems, *Lett. Nuovo Cimento* **25** (1979), 263–265.
8. D. V. CHUDNOVSKY, Topological and algebraic structure of linear problems associated with completely integrable systems, *in* "Proceedings of the XI Conference in Group Theoretical Methods in Physics, Instanbul," Lecture Notes in Physics, Vol. 180, pp. 65–90, Springer-Verlag, Berlin/New York, 1983.
9. D. V. CHUDNOVSKY AND G. V. CHUDNOVSKY, Laws of composition of Bäcklund transformations and the universal form of completely integrable systems in dimensions two and three, *Proc. Nat. Acad. Sci.* **80** (1983), 1774–1777.
10. G. V. CHUDNOVSKY, Rational approximations to linear forms of exponentials and binomials, *Proc. Nat. Acad. Sci.* **80** (1983), 3138–3141.
11. G. V. CHUDNOVSKY, Measures of irrationality, transcendence and algebraic independence, *in* "Journées Arithmétiques 1980" (J. V. Armitage, Ed.), pp. 11–82, Cambridge Univ. Press, London, 1982.
12. G. DARBOUX, "Leçons sur la théorie générale des surfaces," Livres IV, Chapter IX, Cauthier–Villars, Paris, 1896; reprinted by Chelsea, New York, 1972.
13. A. EDREI, Convergence of the complete Padé tables of trigonometric functions, *J. Approx. Theory* **15** (1975), 278–293.
14. J. D. FAY, Theta functions on Riemann surfaces, *in* "Lecture Notes in Mathematics," Vol. 352, Springer-Verlag, Berlin/New York, 1973.
15. R. C. GUNNING, Some curves in Abelian varieties, *Invent. Math.* **66** (1982), 377–389.
16. E. L. INCE, "Ordinary Differential Equations," Dover, New York, 1944.
17. V. G. KAC, D. A. KAZHDAN, J. LEPOWSKY, AND R. L. WILSON, Realization of the basic representations of the Euclidean Lie algebras, *Advan. in Math.* **42** (1981), 83–112.
18. E. R. KOLCHIN, Rational approximation to solutions of algebraic differential equations, *Proc. Amer. Math. Soc.* **10** (1959), 238–244.
19. E. R. KOLCHIN, Differential algebra and algebraic groups, *in* "Pure and Applied Mathematics," Vol. 54, Academic Press, New York/London, 1973.
20. M. G. KREIN, On the continuous analog of the Christoffel formula in the theory of orthogonal polynomials, *Dokl. Acad. Nauk USSR* **113** (1957), 970–973. [Russian]
21. S. LANG, "Diophantine Geometry," Interscience, New York, 1962.
22. S. LANG, Report on Diophantine approximations, *Bull. Soc. Math. France* **93** (1965), 177–192.
23. J. LEPOWSKI AND R. L. WILSON, Construction of the affine Lie algebra $A_1^{(1)}$, *Comm. Math. Phys.* **62** (1978), 43–53.
24. K. MAHLER, "Lectures on Diophantine Approximations," Notre Dame, Indiana, 1961.
25. C. F. OSGOOD, On the effective "diophantine approximation" of algebraic functions over fields of arbitrary characteristic and applications to differential equations, *Proc. Konink. Nederl. Akad. Van. Wetens. Ser. A* **78** (1975), 105–119.
26. C. F. OSGOOD, Concerning a possible "Thue–Siegel–Roth theorem" for algebraic differential equations, *in* "Number Theory and Algebra," pp. 223–234. Academic Press, New York/London, 1977.
27. C. F. OSGOOD, An effective lower bound on the "diophantine" approximation of

algebraic function by rational functions, II, "Contribution to Algebra," pp. 321–322, collection of papers dedicated to Ellis Kolchin, Academic Press, New York/London, 1977.

28. C. F. OSGOOD, Effective bounds on the diophantine approximation of algebraic functions and Nevanlinna theory, *in* "Proceedings of the Number Theory Seminar," Lecture Notes in Mathematics, Springer-Verlag, Berlin/New York, 1984.
29. E. PICARD, "Traite d'Analyse," Vol. III, Chap. 17, Gauthier–Villars, Paris, 1898, 1908, 1928; reprinted as "Analogies Entre la Theorie des Equations Differentielles Lineares et la Théorie des Équations Algébriques," Gauthier–Villars, Paris, 1936.
30. N. RATLIFF, The Thue–Siegel–Roth–Schmidt theorem for algebraic functions, *J. Numer Theory* **10** (1978), 99–126.
31. D. RIDOUT, Rational approximations to algebraic numbers, *Mathematika* **4** (1957), 125–131.
32. K. F. ROTH, Rational approximations to algebraic numbers, *Mathematica* **2** (1955), 1–20; Corrigendum, p. 168.
33. W. M. SCHMIDT, Contributions to diophantine approximations in Fields of Series, *Monatsh. Math.* **87** (1979), 145–165.
34. W. M. SCHMIDT, Diophantine approximation, *in* "Lecture Notes in Mathematics," Vol. 785, Springer-Verlag, Berlin/New York, 1980.
35. C. L. SIEGËL, Uber einige Anwendungen diophantischer Approximationen, *Abh. Preuss. Akad. Wiss.* **1** (1929).
36. S. UCHIYAMA, Rational approximations to algebraic functions, *J. Fac. Sci. Hokkaido Univ. Ser.* **15** (1961), 173–192.
37. J. P. S. KUNG (Ed.), Young tableux in Combinatorics, *in* "Invariant Theory and Algebra. An Anthology of Recent Work," Academic Press, New York/London, 1982.